A Framework for Benchmarking Open Government Data Efforts

Djoko Sigit Sayogo
Center for Technology in Government
University at Albany
dsayogo@ctg.albany.edu

Theresa A. Pardo
Center for Technology in Government
University at Albany
tpardo@ctg.albany.edu

Meghan Cook
Center for Technology in Government
University at Albany
mcook@ctg,albany.edu

Abstract

This paper presents a preliminary exploration on the status of open government data worldwide as well as in-depth evaluation of selected open government data portals. Using web content analysis of the open government data portals from 35 countries, this study outlines the progress of open government data efforts at the national government level. This paper also conducted in-depth evaluation of selected cases to justify the application of a proposed framework for understanding the status of open government data initiatives. This paper suggest that findings of this exploration offer a new-level of understanding of the depth, breath, and impact of current open government data efforts. The review results also point to the different stages of open government data portal development in term of data content, data manipulation capability and participatory and engagement capability. This finding suggests that development of open government portal follows an incremental approach similar to those of e-government development stages in general. Subsequently, this paper offers several observations in terms of policy and practical implication of open government data portal development drawn from the application of the proposed framework.

1. Introduction

Government initiatives to pursue openness by sharing data collected by government with the general public have resurfaced and received widespread attention worldwide in the last 5 years. This effort that was initially inspired by Obama's administration in 2009 claiming that openness will "ensure the public trust and establish a system of transparency, public participation, and collaboration." The initiative to publish government data on publicly accessible data portal began with the launch of data.gov in May 2009 by the U.S. government, followed by the U.K. in September 2009 and New Zealand's beta version portal in November 2009. The publishing data that is collected and stored by government openly to the general public is, herein coined as open government data or abbreviated as OGD.

Open government and open data have been associated with various potential benefits, including the enhancement of social, political, and economic values. For instance, in 2009, President Obama claimed that "openness will strengthen our democracy and promote efficiency and effectiveness in government." Countries committed to open government and open data believe that openness will, among other things, promote transparency, fight corruption and energize civic engagement. Hillary Clinton, during her remarks at the Open Government Partnership (OGP) opening in Brazil in 2012, even went as far as claiming that open data will create a new division among nations based on their openness level, differentiating open versus close nations.

Countries such as Denmark, Spain, U.S. or U.K., have explicitly geared their openness strategies by correlating openness with innovation, new businesses and economic benefits [8]. Robinson, Yu, Zeller and Felten [13] strongly argue that open data enables the general public and private entities to engage in collective learning and garner new public innovations from the data. It is hypothesized that countries that are more open in their economic activities and more open in building economic opportunities through engagement with the general public generate higher wealth and affluence. Thus, in spite of currently lack of empirical justification, the proponents strongly connect open government and open data to economic value both internally for the government agency and externally for the general public.

The benefits of open data and open government are increasingly recognized by a range of national and international organizations, as well as political figures [4]. The UN open data handbook lists various claims of benefits including cost efficiency for government. David Cameron, U.K. Prime Minister, explicitly specified in a letter U.K. government department in May 2010 that the objective of open data in the U.K. was to "reduce the

978-1-4799-2504-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HICSS.2014.240

deficit and deliver better value for money in public spending; and to realize significant economic benefits by enabling businesses and non-profit organizations to build innovative applications and websites using public data".

Government attention and interest in providing access to and use of government data is recently magnified through inter-government partnerships in open government called the Open Government Partnership (OGP) launched in September 2011. One of the eligibility criteria for joining OGP is for the country to ensure access to information, including access to government data as well as laws governing access. We argue that participation in the OGP adds supplementary motivation to develop and implement OGD efforts. Looking at similar occurrences such as the United Nations e-Government Survey, it is safe to assume that increased visibility and peer group pressure imposed through participation in OGP motivates rapid progression toward open government data adoption especially at the national level both in developed and developing countries.

Rapid adoptions of OGD and open government concepts worldwide raised a concern over the ability of the national government to actually keep up with their promise in progressing toward OGD efforts. Learning from the lessons from empirical investigations in numerous studies on e-government implementation, progress of information technology application in government has shown to follow an incremental process [10] in part due to the various barriers to adoption and implementation efforts [3, 6]. Likewise, various factors, from institutional to technical, seem to affect the development and implementation of OGD portal at the national level [9]. Thus, it is sensible to argue that different nations have different capabilities in developing and implementing their OGD efforts.

On the other hand, there are very few studies or resources at present that evaluate the progress of OGD efforts in terms of meeting a set of specified goals and then describing and comparing efforts across countries. This paper presents a preliminary exploration of the status of OGD efforts and subsequently proposes a framework to facilitate evaluation of OGD status. Using web content analysis of OGD portals from 35 countries, this study aims to outline the progress of government open data efforts at the national government level. This study addresses the question of how the disclosure of data is structured, organized, and published, in terms of scope, complexity and engagement capability. The findings of this review offer a new-level of understanding of the depth, breath, and impact of current OGD portals. The findings from the reviews as well as the proposed framework are useful for 1) understanding the progress of OGD efforts in terms of meeting a set of specified goals and 2) providing benchmarks to assist open data planners in designing OGD initiatives in terms of scope, complexity, and engagement capability.

The rest of the paper proceeds as follows. Section two presents an overview of current literature in open government data in terms of definition, timeline and principles, as well as the concept for benchmarking open government efforts. Section three outlines the research approach used in this paper. Section four describes the results and findings from the analysis of the OGD portal from the 35 countries. This section also presents review of the portals in term of two capabilities, data manipulation and community engagement. Finally, section five provides concluding remarks.

2. Literature Review

2.1. The Definition of Open Government Data (OGD)

The idea of open government data re-emerges with increased attention to open government and open data initiatives and rapid development of information technology promoting interactivity. Thus, both open government and open data initiatives are not entirely a new concept. Extant literature pointed to the historical existence of both open government and open data. The idea of open government has an older origin, dated back to the early years following the World War II at the time of great depression in the U.S. There were notable efforts to increase pressure toward openness and application of the right to know principle in the government during that period [12]. The term "open data" is not as old as open government but was initially recognized in the natural science field. The term was used to indicate basic and untreated scientific data. The first recognition on the used of "open data" to refer to a policy context, defining a scientific policy for a research project, was in the 1970s, during an international collaboration project in NASA [16].

At the conceptual level, the idea of openness has not significantly changed. For instance, the definition of open data in the current context still includes several important elements characterizing the preceding definition such as accessibility, availability, re-usability, re-distribution and participation [5] in combination with the emphasis on information technology factors such as machine-processable, non-proprietary and license-free [1]. On the other hand, current openness as a concept is wider due to the emphasis on information technology usage and innovation.

On the other hand, despite the long history of both open government and open data, the definition of open government data is ambiguous. Some studies argue that

open government data represents a convergence between open government and open data. However, the term "open government data" is itself ambiguous [16], leading to questions such as what is government data. As consequences, there are various conceptualizations of open government data. Some studies provide strict definitions such as the definition proposed by the Open Knowledge Foundation which defines OGD as "data produced or commissioned by government or government controlled entities." This definition restricts the scope of government data to include only data produced or commissioned by government. This definition omits data that is reported by private entities and is collected and stored by government agencies. Other studies expand the conceptualization of OGD by positing the principles of open data and not an exact definition. For instance, the Sunlight Foundation's open government working group posit that government data is open if the way it is structured and published met eight principles. The Sunlight Foundation revised the eight principles by adding two additional principles to define open government data. A more specific definition was given by Geiger and Lucke [7] asserting that open government data is "all stored data of the public sector which could be made accessible by government in a public interest without any restrictions for usage and distribution."

This paper conceptualizes the definition of open government data (OGD) loosely to indicate open publication of data collected and stored by government agencies which is consented by laws to be made accessible to the general public through a single data portal. This conceptualization recognizes three elements constituting an OGD. First, an OGD publishes all data that is collected and stored by government agencies. Second, an OGD only publishes data that is consented by law, including privacy, confidentiality and security principles. Finally, an OGD allows general public to access data through a single data portal.

2.2. Open Data Principles as Indicators of Progress on Opening Government Data

In 2010, Sunlight Foundation published ten principles governing open data, especially targeted to open government data. These ten principles are based on an updated eight principles developed by the Sunlight Foundation's Open Government working group in 2007. Sunlight Foundation brought together a group of 30 open government advocates in 2007 to discuss open government and open data. The group, led by Carl Malamud and Tim O'Rielly, generated eight principles of open government data. These principles were expanded by the Sunlight Foundation [14] to include two additional principles governing the permanence of data and cost of using the data. The ten principles and the indicators for measuring each principle are presented in Table 1. The Sunlight Foundation also asserts that "each principle exists along a continuum of openness", although they did not provide clarification of how each principle is used in a continuum of openness.

Table 1. Ten Principles of Open Data

Principle	Indicator
Completeness [a,b]	• Datasets should be as complete as possible and includes metadata • Datasets do not compromise with federal laws on privacy, security of privilege
Primacy [a,b]	• Datasets should be primary source and include original information • Includes details of collection process
Timeliness [a,b]	• Available in timely fashion and be released as quickly as possible • Release priority depends on the time sensitivity of data
Ease of Physical and Electronic Access [a,b]	• Datasets should be accessible and easy to obtained by any means • Pay attention to the "findability" of datasets
Machine Readability [a,b]	• Information should be stored in widely-used file formats and that enable machine processing • Documentation related to the format and how to use it is provided.
Non-discrimination [a,b]	• Any person can access the data at any time and without having to identify themselves or provide any justification for doing so
Use of Commonly Owned Standards [a,b]	• Used freely available alternative formats • There is no software license needed for accessed and used.
Licensing [a,b]	• Clear labeling public information as a work of the government • Not subject: copyright, patent, trademark or trade secret regulation
Permanence [a]	• Online stickiness - information is available online in perpetuity
Usage Costs [a]	• No cost for accessing the data, or cost is as minimum as possible

a) Sunlight Foundation principles (2010)
b) Open Government working group principles (2007)

The Open Data principles framework published by the Sunlight Foundation primarily govern the presentation and storage of data. This framework is proposed to ensure optimum accessibility, availability and re-usability.

This framework suggests that presentation of complete, primary, and timely data are necessary to ensure optimum availability. The administrator should ensure ease of access, non-discrimination, free license and minimum usage costs to increase accessibility of the data. Data should be formatted to conform to machine-readable format, use common standards and be available perpetually online to improve usability. On the other hand, the framework is less useful to measure online capabilities and features available in OGD portals to facilitate users to use the data. The next section describes a proposed framework of data manipulation and engagement capability for benchmarking OGD status.

2.3. Data Manipulation and Engagement Capability for Benchmarking OGD status

This paper defines online capabilities and features as availability of features capable of facilitating data manipulation and user's engagement such as visualization tool, online rating, forum, etc. This paper proposes a simple concept of structure (figure 1) to review the progress of open data portal by recognizing data representation capability as well as online capabilities/features to facilitate users' experience. In this sense, this proposed framework complements the framework based on open data principles.

		Data Manipulation Capability		
		No Features	Limited	Advance
Engagement Capability	No Features			
	Participative			
	Collaborative			

Fig 1. The Proposed Framework for Benchmarking

We conceptualize a framework consisting of two elements, a) data manipulation and b) engagement capability. These elements can be used independently or in combination. In addition, this study recognizes that the ability of government to manage their open data portals evolved along with time and experience of government technology use. As evidenced from a number of studies of e-government development, typical reform in government policies does not occur through revolutionary transformation but rather, evolves gradually [15]. As such, government's ability to provide online features promoting data manipulation, interactivity and community engagement is also affected by time and experience.

The adoption of a more sophisticated design that could promote openness through data manipulation and engagement develops gradually as governments gain more experience. Consequently, we argue that there are continuums of online capability development to support data manipulation and engagement applicable to OGD portal. The proposed framework is comprised of a continuum of none to limited to advance online capabilities for data manipulation and engagement. The next two sub-sections describe the data manipulation and engagement capability in details.

2.3.1. Data Manipulation Capability. This capability refers to the online features and data presentation that facilitate users to manipulate the data. On December 17, 2004, the Office of Management and Budget (OMB) of the Executive Office of the President of the U.S., release memorandum specifying policies for federal agency public websites. The memorandum specified the objectives of publishing government data was "permitting users to aggregate, disaggregate, or otherwise manipulate and analyze the data to meet their needs [11]." Consequently, publication of government data through open data portal should enable data manipulation.

The application of the data manipulation capability continuums is based on the combination of the two elements, a) the online features capable of facilitating data manipulation and b) the data representation in the OGD portal. The specification of these continuums is as follow:

- **No Features**. This capability refers to the portal without features to enable manipulation of data such as sorting, amassing, and others, and presents data that is not complete, in not in primary source, and/or not timely.
- **Limited**. This capability refers to the portal that presents data allowing for limited analysis and provides limited online features to enable users to manipulate data. For example, portal that publishes data only in xls or xlsx format and only allows users to sort data.
- **Advanced**. This capability refers to the portal providing advanced data manipulation capabilities, such as amassing different data and publishing data in the format that enable users to conduct data analysis.

2.3.2. Engagement Capability. This refers to the online features of an open data portal that enable engagement with and/or among users. One of the major pillars of open government as promoted by President Obama in 2009 is collaboration, defined as "use innovative tools, methods, and systems to cooperate…across all levels of government and with nonprofit organizations, businesses, and individuals in the private sector" Consequently, providing engagement capability is important to facilitate user's interactions necessary to foster innovation, just like the idea of crowdsourcing.

The application of this framework is based on availability of features to solicit participation and engagement of users through an OGD portal. The

applicability of the framework also considers participation and collaboration. Participation refers to the extent to which users can participate in the governance of the OGD portal such as by suggesting datasets or voting for performance of the portal. Collaboration implies the extent to which the OGD portal enables or facilitates cooperation and collaboration among government at all levels, nonprofit organizations, businesses and individuals. The elements are divided into three:

- **No Features**. The OGD portal does not provide tools and or systems for users to engage with government and/or other users.
- **Participative**. The OGD portal provides tools and systems to enable users to participate in the governance of the portal such as voting or rating but the engagement with other users is limited or mediated by the administrator.
- **Collaborative**. The OGD portal provides tools, methods and systems to enable users to innovatively engage and collaborate with other users be it government at any level, nonprofit organizations, businesses or other enthusiasts

3. Methodology

This paper employs two approaches in reviewing the status of OGD portals, a) overall assessment of the status of OGD portals of 35 national governments and b) review of eight selected cases of OGD portals to showcase the applicability of the proposed framework for measuring online capability.

a. Overall assessment of the status of OGD portals.

The first approach is reviewing the status and progress of the national government's OGD portal worldwide. The review of these portals focused on the data representation and availability based on the open data principles propagated by the Sunlight Foundation. The first step of this review was identifying and compiling the existing OGD portals. This identification and compilation process followed two steps. First is initial identification of the OGD portals based on the listing on the U.S. data.gov portal. The open data sites menu in the U.S. data.gov listed 41 countries as having national level open data portals. Second is validating and verifying the OGD portals identified from the U.S. data.gov.

The process to verify and validate the viability of the listing in U.S. data.gov followed three steps:

1) We inputted a keyword consisting of the name of the countries listed in the U.S. data gov and added "open data" into general search engine, Google,
2) We compared the list in the U.S. data.gov with the list available at other sources such as datacatalogs.org,
3) We opened the identified portal to examine whether the sites are in working condition.

After conducting the verification process, we reviewed OGD portals from 35 countries. This review was conducted in two waves, first wave was in December 2012 and second wave was in February 2013. The data presented in this paper reflected the status of the OGD portals after the second wave review.

b. Review of selected cases of OGD portal.

The second approach was conducting detailed assessment on set of eight OGD portals. The review of these portals was conducted based on the framework for measuring online capability. The objective of this second phase review is to showcase the usefulness of the proposed framework for evaluating the different degree to which OGD portals present online capability supporting data manipulation and engagement features. The eight OGD portals were selected purposively to represent a) open data efforts by government in developing and develop countries and b) open data efforts by multi-national organizations. The selection of the eight portals also followed convenience sampling logic by primarily selected countries in the continent of Africa. The selected countries and multi-national organizations are: Morocco, United Arab Emirates, Ghana, Kenya, India, U.S., Worldbank and Africa Development Bank (AfDB).

4. Results

This section presents the results of our assessment in three sections. The first section represents an overall overview of OGD portal status and progress worldwide in terms of data representation. The second section focuses on the overview of the OGD portal status and progress in term of participatory and engagement capability. The third presents the review of online capabilities and features of selected OGD to facilitate data manipulation and engagement, as well as demonstrating the use of the proposed framework.

4.1. General Status and Compliance to the Open Data Principles

The findings of the review on the general status of the OGD portals is provided in figure 2. The findings indicate that not all countries have an open data portal analogous to the U.S. or UK. The majority of OGD portals, around 92%, developed their OGD portal independent of the national government portal, while only three countries integrated the open data portal with the National Government portal. We also found only two countries, Ireland and Estonia that have portals as an expansion of their National Statistical Bureau sites at the time of the review. The findings also show that a majority of the OGD portals (95%) are spearheaded by a government agency, particularly a federal government

agency, and only two out of the 35 countries, Russia and Peru, have their open data initiated by non-government.

Our review indicates that the majority of OGD portals complied with the open data principles in terms of providing granular data, accessibility, and share-ability of data (Figure 3). We found that the majority (89%) of the 35 OGD portals provide granular level of data and raw data, while only 8% provide aggregates or modified data. More than half of the OGD portals reviewed (66%) offer data in machine readable format thus increasing the likelihood users can share and manipulate data.

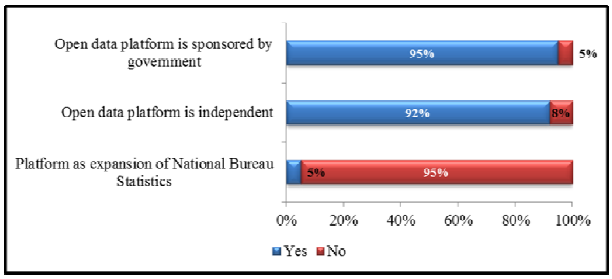

Fig 2. General Status of OGD Portals

A greater part of the OGD portals (80%) adopt an Open License agreement that could increase the accessibility of data. Accessibility is also supported by the use of languages other than the country's own national language by a little more than half (57%) of the OGD portals.

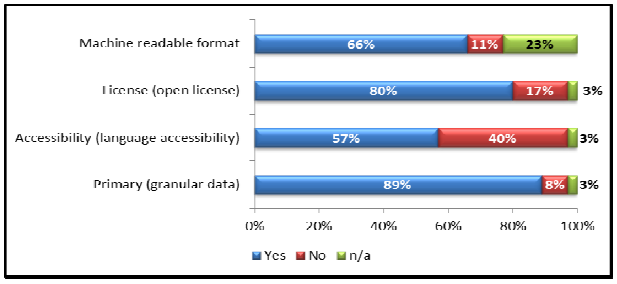

Fig 3. Compliance to the Open Data Principles

4.2. Participatory, Engagement Capability and Share-ability of Data in the OGD Portals

This section focuses on OGD portal status and progress in term of participatory and engagement capability, as well as the share-ability of data in the portal. This section consists of two sub-sections.

4.2.1. Participatory and Engagement Capability. This sub-section analyzes the extent to which effort was made in each OGD portal available to stimulate participation and engagement, as one of the primary characteristics of Open Government. We found five mediums to solicit participation and engagement in the OGD portal (Figure 4). These mediums are:

a. Engagement through data request and ideas. One way of engaging users is through the request for datasets. The idea is to foster involvement from the public to enrich the repository of the datasets that reflects the public needs. The results show that 49% of OGD portals provide features facilitating the users to send request for dataset suggestions, while 51% of portals have not yet providing this feature.

b. Engagement through sharing of Apps and ideas. Based on the review of the OGD portal, the results show that 51% of the portals hosted online or mobile applications (shortened as apps) in their portal with only 49% have not yet providing this feature. These apps are generated by government, citizens and private entities. For instance, the US open data portal features 1,264 apps built by government and 236 apps built by general public (citizen and private entity).

c. Participation through discussion. The other medium facilitating participation and engagement is through lively discussion in the portal. The findings indicate that 57% of the OGD portals provides feature to foster discussion, while the 43% did not provide any tools to encourage discussion in their open data portal. There are various forms of discussion features, namely: community, forum, feedback and comments. Different discussion features provides different richness of discussion mediums. For instance, feedback and comment facilitate one-way engagement in which the users could send their comment or feedback and then waiting for the administrator to reply. Community and forum, on the other hand, provides more specified and livelier space for discussion.

d. Reach and engagement. We define reach and engagement as the effort to spread and increase the usage of the open data portal. Majority (71%) of the OGD portals employ some kind of reach and engagement features. The majority of the portals (80%) that provide this feature are commonly use social media to support their reach and engagement effort. The 20% of the countries that did not use social media as engagement tool (20%) use alert tool such as RSS feed to support their engagement effort. The 29% of the OGD portals that did not use either social media or RSS feed relies on mass media or not providing engagement tools at all.

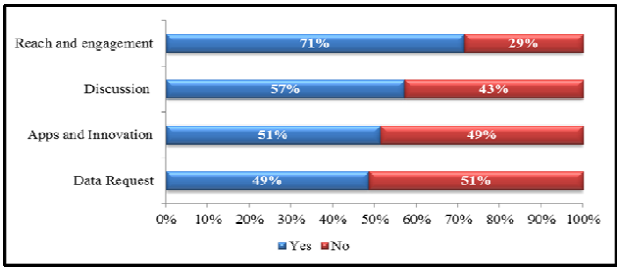

Fig 4. Engagement and Participation Effort

e. Engagement through Involvement in Measuring Performance of the OGD. This feature refers to the measurement of the usefulness and quality of the data provided in the open data portal. This measurement represents another indicator of public engagement in the OGD portals through rating or ranking the data provided by government agencies in the portal. Majority of the OGD portals (80%) did not provide feature to engage users in measuring performance of the data while only 20% provide this tool. Of these 20%, all of them use rating or scoring system as metrics to measure the quality and usefulness of data.

4.2.2. Accessibility and Share-ability of OGD Portals. This section evaluates the extent to which the open data platform offers functionality for ease of analysis, accessibility, share-ability, and linked data (figure 5). The review indicates that 66% of the existing open data portals provide the ability to manipulate the use of data, from the ability to search to the ability to mash-ups data. The open data portals (34%) that did not provide ability to manipulate the data only provide a list of available data. Another feature to support ease of analysis of the data is visualization. Only 20% of the existing OGD portal provides visualization features, from limited basic charts and maps to advanced charting and visualization. Currently only the U.S. OGD portal provides advanced visualization and charting for the users.

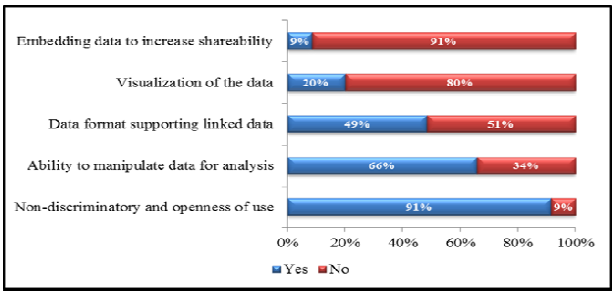

Fig 5. Accessibility, Share-ability and Linked Data Capability

In terms of share-ability, the ability to share the data, tools and content in another site, only 9% of the existing OGD portals provide features to embed datasets and data tables to other sites while 91% are not yet providing this feature. The share-ability in term of non-discriminatory and openness of use shows the opposite result with only 9% of the existing data portals did not provide support for openness. For instance, the open data portal of Uruguay requires users to register to be able to download the data.

The review of the OGD portals also indicates that 49% of the existing data portal provides data in formats that support linked data while 51% provides data that is not support linked data. The portals that support linked data (49%) publish data in RDF, JSON, XML, and API. The portals that did not support linked data (51%) only provide data in proprietary format.

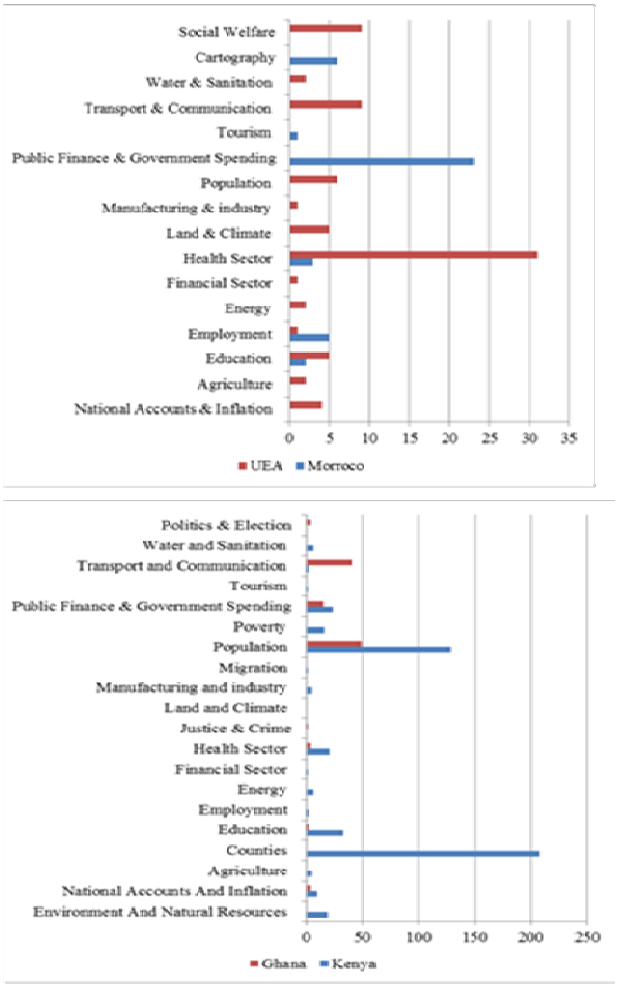

Fig 6. Data Content from Four Countries

4.3. Data Manipulation and Engagement Capability of Selected OGD Portals

In this section, we present eight OGD portals with six portals from the national government and two portals from international organizations. We review the portals based on the proposed framework outlined in section 2.3 focusing on data manipulation and engagement capabilities of the OGD portals to showcase the applicability of the framework.

To begin the review and before we apply the framework, we focus on the data content of the OGD portals from six national governments in Africa continent that offer limited engagement capability (Table 2). We focus on Morocco, United Arab Emirates, Ghana and Kenya. The objective is to understand the focus of data published in the OGD portals to understand if the focus differs among countries. We found each country focused on different types of data in

their portals (Figure 6). The United Emirates of Arab (UEA) published a large number of data pertinent to the health sector. Morocco focused more on public finance and government spending. Ghana provides a large number of data on communication, population-related data, and transportation such as the number of accidents. Kenya focused on providing data related to demographic data of counties such as urbanization.

The review of the four OGD portals also points to the significance of reviewing the growth of data provided in the portal as indication of progress and performance. For instance, Kenya's OGD portal indicates approximately 37% increase of datasets availability from November 2011 to February 2013. In February 2013, there were 534 datasets available in the portal, an increase of 104 datasets from mid-2012 and an increase of 144 datasets from November 2011.

Table 2. Data Content for 4 Selected Countries

Data Category	MAR	UEA	GHA	KEN
Number of datasets & growth rate*)	38 datasets	78 datasets	122 datasets	534 data (Feb 2013) & 144 data (Nov2011)
Format of data	Primarily in xls format	xml, xls, pdf & word	Primarily in xls format	Varies e.g., xml, json, html, csv, xls, etc.
Majority of Data	Public Finance (23 data)	Health Sector (31 data)	Demo-graphy & Pop. (29 data)	Demo-graphy of counties (208 data)
Most viewed data	n/a	n/a	Cost of student per capita (1,085 views)	Poverty rate by district (19,263 views)

Our subsequent reviews focus on the availability of data manipulation and engagement capability. The two countries, Morocco and UAE, did not provide online capability to manipulate data. The OGD portal of these two countries only listed the available data. The OGD portal of Ghana provides limited data manipulation capability. The OGD portal of Ghana provides capability in term of data filtering and sorting and does not provide any visualization capability.

The OGD portal of Kenya provides more advanced data manipulation capability than Ghana, in term of data visualization and data filtering but does not provide data mash-up capability. Using the visualization and filtering capability, the users can create graphic or map visualization based on particular types of data or subsets of data. For example: map visualizations of health facilities in Kenya (see figure 7).

Our review of the two international organizations, AfDB and Worldbank, indicates that both offer advanced capability for manipulating data such as: data mashups and maps, dashboard, data query, and data filtering.

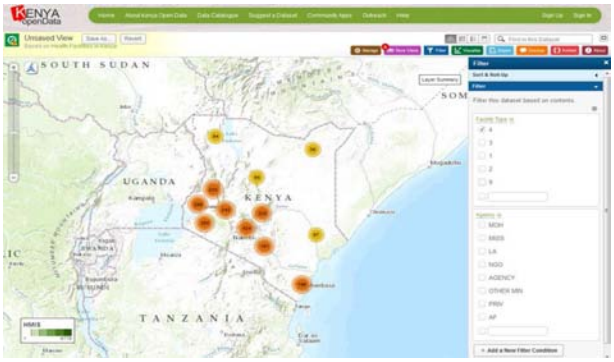

Fig 7. Snapshot of Map Vizualization of Health Facilities in Kenya

AfDB provides access to maps, country profiles/dashboards, interactive data queries, and the ability to mashup data. The AfDB portal provides five important features, namely: maps, dashboard, data query and data by topic. Similar to AfDB, the World Bank's databank provides advanced data manipulation capability in term of visualization (graphs and maps), country profiles/dashboards, interactive data queries, and the ability to mashup data. The datasets are categorized based on country, topic, data indicators, data catalogue, and microdata. A summary of data manipulation capability of the four countries and two international organizations is provided in Table 3.

Table 3. Data Manipulation Capability

	Data Format (complete, primacy, timelines, machine-readable, etc)	**Online Features** (searching, sorting, amassing, etc)
No	Morocco, UEA	n/a
Limited	Ghana	Ghana
Advance	Kenya, AfDB, Worldbank	Kenya, AfDB, Worldbank

In terms of engagement capability, the OGD data portals of Morocco and UAE did not provide any features to facilitate engagement with users. Their portal simply focused on publishing data, similar to the catalogue stage in Layne & Lee [10] model or the presence stage of Gartner's model [2], or other e-government maturity model. The OGD portal of AfDB and Worldbank provide advanced capability for data manipulation. However, the OGD portal of AfDB and Worlbank did not provide any means to facilitate participation except through a "contact us" menu where users can send their feedback.

The summary of data manipulation and engagement capability of the six countries and two international organizations is provided in Table 4. The OGD portals of Ghana and Kenya provide limited capability for

user's participation. The OGD of Ghana provides three types of user ratings, namely: usability, quality and accessibility. The portal also provides visitor statistics, a visualization of visitor activities in the portal such as number of downloads. In addition, the Ghana OGD portal provides features for datasets suggestion. The OGD portal of Kenya also provides user rating as a way to promote user's participation. There are four types of rating system which are: most relevant, most accessed, highest rated and most comments. The most relevant rating is equivalent to the most views datasets, indicating the number of views but not necessarily the number of accessed/download. The OGD portal of Kenya also features suggested datasets. There are three statuses of suggested datasets, which are: open, approved and reject. Open indicates that government officials have not responded to the request. The status changes to approved or reject once the suggestion undergoes review by a government official.

Table 4. Data Manipulation and Engagement

		Data Manipulation Capability		
		No Features	Limited	Advance
Engagement Capability	No Features	Only publish data in format that allow for limited analysis [eg. Morocco, UEA]		Provide data manipulation capability in the portal but practically no participation [eg. AfDB, Worldbank]
	Participative		Publish data in format complying with open data principles and provide limited participation [eg. Ghana]	Provide data manipulation capability in the portal but very limited participation [eg. Kenya]
	Collaborative		Provide limited data manipulation capability and allow for community formation around data topic [eg. India]	Provide data manipulation capability and allow for community formation around data topic [eg. US]

Examples of OGD portals providing advanced engagement capability are India and the U.S. Both portals provide features such as: blogs to facilitate the sharing of resources and knowledge, a forum to facilitate discussion on applications and technology, and community. Community engagement refers to the features that enable connectivity, interaction and sharing of data, views and insights through public discussions and open forums to stimulate public innovation. Each community aims at a particular topic of interest, such as: energy, health, green supply chain and other. The OGD portal of India supports three communities, namely: developers, health and agriculture. The U.S. OGD portal has the most advanced capabilities in terms of both data manipulation features and support for community-based engagement. There are currently 17 different communities supporting 17 different topics in their portal. Each community has different activities and each community can have many different activities within their community. However, there are three core activities that each community the U.S. OGD portals must support, namely: data/tools, apps and forum/let's talk. The U.S. government also uses challenges and prizes to stimulate wider engagement and participation.

5. Concluding Remark

Using web content analysis of OGD portals from 35 countries, this study outlines the progress of OGD efforts at the national government level and proposes a framework for understanding the status of OGD initiatives. Open government data (OGD) has become the vanguard of the modern open government movement. Implementation of OGD becomes political conventions of various political leaders worldwide. The increased attention and rapid adoption, on the other hand, raises concerns over the progress of OGD efforts. This paper presents a preliminary exploration of the status of OGD worldwide as well as in-depth review of selected OGD portals to demonstrate the application of data manipulation and engagement capability. The findings from the reviews as well as the proposed framework are useful for 1) understanding progress of OGD efforts in terms of meeting a set of specified goals and 2) providing benchmarks to assist open data planners in designing OGD initiatives in terms of scope, complexity, and engagement capability.

The findings of this review offer a new understanding of the scale and development of current OGD portal initiatives. The review results also point to the different stages of OGD portals development in term of data content, data manipulation capability and participatory and engagement capability. This finding suggests that in general, OGD portals development follows an incremental approach similar to those of e-government development stages.

This paper then concludes with several observations in terms of policy and practical implication of OGD portal development drawn from the application of the proposed framework to eight selected cases.

a. Attention should be given to the correlation between focus of data being published and the public need. The review results indicate that each OGD portal has a different focus on data being published. Morocco focuses on public finance, UEA on health sector, Kenya on urbanization and Ghana on transportation. On the other hand, the focus of data being published does not correlate with the data that most viewed by users. Ghana focuses on communication and transportation while their most viewed data is cost per student per capita. There are two plausible explanations for this condition. First, there is misalignment between data that the public need and the government decision on publishing the data. Anecdotally, government might not take the public need into consideration when deciding the publication of data. Second, there is an issue with data availability. The government is, in rush on publishing

data to comply with the pressure to show progress with their OGD efforts, published any available data through their OGD portals. Consequently, this assertion calls for further research to ascertain the process of publishing data in the OGD portal and the extent to which government consider public need in publishing their data. As well as research to understand the perception of government agencies to support open data initiative promoted by the central government.

b. Governments need to invest significant time to allow agencies to prepare data for publication. The issue of data availability as argued above indicates that central governments need to give agencies time to prepare data to be published in the OGD portals. The process of requesting for data from agencies need to begin ahead of time before the development of the portal itself. For instance, Kenya started requesting data from the agencies in 2009 for launching the beta version of OGD portal in 2010.

c. The format of data to be published influences the extent to which users can innovate or use the data. Government need to consider which format to publish their data. There are many elements to consider the format of data to be published, such as: the format of currently available data, availability of resources and technical capability and others. This anecdotal assertion necessitates a further research to provide empirical justification on the correlation between the types of data format and public innovation.

d. The types of user engagement provided in the OGD portal might correlate to the level of user's engagement. The review indicates different types of user's engagement and participation provided in the OGD portal, from user ratings to community-based engagement. The different types of participatory tool might affect the extent to which users can participate, as well as using the portal to engage in collaborative innovation among users.

Finally, the results presented in this paper were based on the content analysis of OGD portal conducted in two waves, in December 2012 and February 2013. OGD portals as any other web portals are dynamic, meaning that the content and capability of the portal are continuously changing and evolving. Consequently, the results presented on this paper should be construed to represent the content and capability of OGD portal at the time of exploration

6. References

[1] 8 Principles of Open Government Data: 2007. *http://www.opengovdata.org/home/8principles.*

[2] Baum, C.H. and Di Maio, A. 2000. *Gartner's Four Phases of E-Government Model*. Technical Report #317292. Gartner Group.

[3] Coursey, D. and Norris, D.F. 2008. Models of E-Government: Are They Correct? An Empirical Assessment. *Public Administration Review*. 68, 3 (2008), 523–536.

[4] Davies, T. 2010. Open Data, Democracy and Public Sector Reform: A Look at Open Government Data Use From data.gov.uk.

[5] Dietrich, D. et al. 2012. Open Data Handboook Documentation: Release 1.0.0.

[6] Ebrahim, Z. and Irani, Z. 2005. E-government adoption: architecture and barriers. *Business Process Management Journal*. 11, 5 (Oct. 2005), 589–611.

[7] Geiger, C.P. and von Lucke, J. 2012. Open Government and (Linked) (Open) (Government) (Data). *Journal of eDemocracy - JeDEM*. 4, 2 (2012), 265–278.

[8] Huijboom, N. and Van den Broek, T. 2011. Open Data: an International Comparison of Strategies. *European Journal of ePractice*. 12, (2011), 1 – 13.

[9] Janssen, M. et al. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*. 29, 4 (2012), 258–268.

[10] Layne, K. and Lee, J. 2001. Developing fully functional E-government: A four stage model. *Government Information Quarterly*. 18, 2 (2001), 122–136.

[11] OMB 2004. *Policies for Federal Agencies Public Websites*. Technical Report #M-05-04. Office of Management and Budget.

[12] Parks, W. 1957. Open Government Principle: Applying the Right to Know Under the Constitution. *The George Washington Law Review*. 26, 1 (1957), 1–22.

[13] Robinson, D. et al. 2009. Government Data and the Invisible Hand. *Yale Journal of Law & Technology*. 11, (2009), 160.

[14] Ten Principles for Opening Up Government Information - Sunlight Foundation: 2010. *http://sunlightfoundation.com/policy/documents/ten-open-data-principles/*. Accessed: 2013-05-31.

[15] West, D.M. 2004. E-Government and the Transformation of Service Delivery and Citizen Attitudes. *Public Administration Review*. 64, 1 (2004), 15–27.

[16] Yu, H.M.-T. 2012. *Designing Software to Shape Open Government Policy*. Princeton University